Computer Science Department

TECHNICAL REPORT

"On Two Conjectures Regarding Eigenvalue Perturbations and a Common Counterexample"

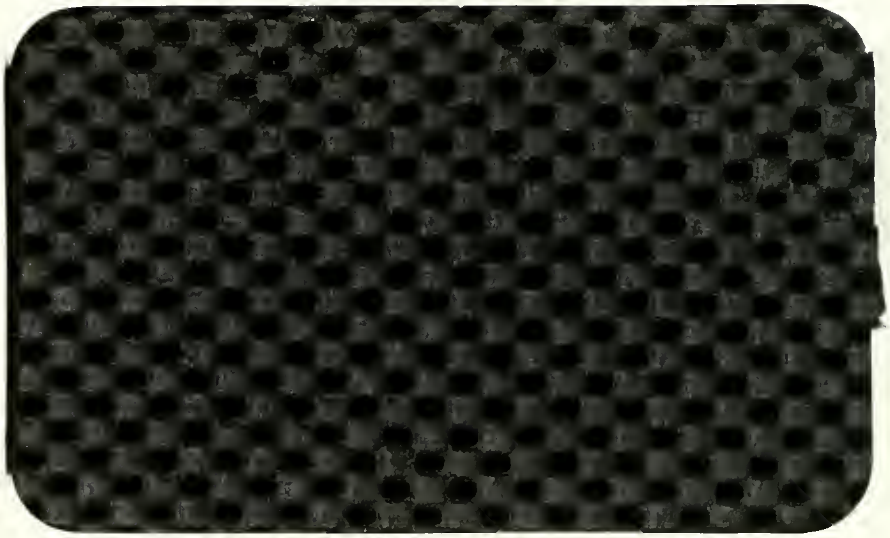
by

James Weldon Demmel

Technical Report #220
May, 1986

NEW YORK UNIVERSITY

Department of Computer Science
Courant Institute of Mathematical Sciences
251 MERCER STREET, NEW YORK, N.Y. 10012



"On Two Conjectures Regarding Eigenvalue Perturbations and a Common Counterexample"
by
James Weldon Demmel

Technical Report #220
May, 1986

On Two Conjectures Regarding Eigenvalue Perturbations and a Common Counterexample

James Weldon Demmel
Courant Institute of Mathematical Sciences
251 Mercer Str.
New York, NY 10012

Abstract

Recently Van Loan and Demmel made conjectures about eigenvalue perturbations. Van Loan's conjecture concerned the smallest perturbation that makes a stable matrix unstable, and Demmel's concerned the smallest perturbations that makes two matrices with disjoint spectra have a common eigenvalue. We show that the truth of either of these conjectures would imply the truth of a third weaker conjecture for which we supply a counterexample.

Recently Van Loan and Demmel made reasonable sounding conjectures about eigenvalue perturbations. Van Loan's conjecture concerned the smallest perturbation that makes a stable matrix unstable, and Demmel's concerned the smallest perturbations that makes two matrices with disjoint spectra have a common eigenvalue. We show that the truth of either of these conjectures would imply the truth of a third weaker and equally reasonable sounding conjecture. We then present a counterexample for this third conjecture which hence also serves as a counterexample to the first two.

A *stable matrix* is a matrix all of whose eigenvalues have negative real parts. Van Loan [Van Loan] recently made the following conjecture about the smallest perturbation of a stable matrix A which makes it unstable ($\|\cdot\|$ denotes the 2-norm):

Conjecture 1: Let A be stable. Let B be the closest unstable matrix to A, i.e. B is unstable and minimizes $\|A-C\|$ over all unstable C. Then B has an eigenvalue on the imaginary axis with the same imaginary part as some eigenvalue of A.

If this conjecture were true, it would lead to a simple computational scheme for computing $\|A-B\|$:

$$\|A-B\| = \min_{\lambda \in \sigma(A)} \sigma_{\min}(A - i \operatorname{Im}\lambda\, I)$$

where $\sigma(A)$ is the spectrum of A and $i=\sqrt{-1}$.

$\mathrm{sep}_\lambda(A,B)$ is the size of the smallest perturbations to A and B which makes them have a common eigenvalue:

$$\mathrm{sep}_\lambda(A,B) \equiv \min_\lambda \max(\sigma_{\min}(A-\lambda I)\ ,\ \sigma_{\min}(B-\lambda I)) \ .$$

Let $\sigma(A)$ denote the set of eigenvalues of A and similarly for $\sigma(B)$. Let $\mathrm{co}(X)$ denote the convex hull in the complex plane of the point set X. Demmel made the following conjecture about the minimizing λ in the definition of sep_λ:

Conjecture 2: The minimizing λ in the definition of sep_λ above lies in the convex hull $\mathrm{co}(\{\sigma(A)\ ,\ \sigma(B)\})$ of the spectra of A and B.

If this conjecture were true, it would greatly limit the region in the λ-plane that had to be searched for the minimizing λ.

In this short note we show if either of these two conjectures were true then a third weaker and equally reasonable sounding conjecture would be true. Then we will present a counterexample to this third conjecture. First we need some notation. Define $S(A,\epsilon)$ as the set of all eigenvalues of all matrices $A+\delta A$ for $\|\delta A\| \le \epsilon$:

$$S(A,\epsilon) = \{\lambda : \det(A+\delta A-\lambda I)=0 \ , \ \|\delta A\| \le \epsilon\} \ .$$

Conjecture 3: Let A be a matrix with a single eigenvalue λ. Then $S(A,\epsilon)$ is convex.

It is easy to see how Conjecture 3 is implied by either of the first two conjectures. First consider Conjecture 1. Note that $S(A,\epsilon)$ is connected, since any component must contain λ. As ϵ increases, $S(A,\epsilon)$ grows from a single point λ for $\epsilon=0$ to larger and larger sets. The value of ϵ for which this set first touches the imaginary axis is the size of the smallest perturbation that makes A unstable. Suppose there is a matrix A which violates conjecture 3. By multiplying A by a complex number ω of absolute value 1 and adding a multiple α of the identity, we can rotate and shift the eigenvalues of A so that A is stable and $S(A,\epsilon)$ makes any angle to the imaginary axis we want. Since $S(A,\epsilon)$ is nonconvex, we can clearly choose ω and α so that $S(\omega A+\alpha I,\epsilon)$ appears as in Figure 1. By varying ω and α slightly from these values, we can clearly make $S(\omega A+\alpha I,\epsilon)$ either first touch the imaginary axis only at a single point above the origin, or at a single point below the origin. Thus we can guarantee that it does not touch directly to the right of the single eigenvalue of $\omega A+\alpha I$. Thus Conjecture 1 is clearly violated. Therefore the truth of Conjecture 1 would imply the truth of Conjecture 3.

Now consider Conjecture 2. Again we proceed by contradiction. If Conjecture 3 were false, we could find an A with a single eigenvalue and an ϵ such that $S(A,\epsilon)$ were nonconvex. Again choose ω and α so that $S(\omega A+\alpha I,\epsilon)$ appears as in Figure 1, with the additional condition that the two points where $S(\omega A+\alpha I,\epsilon)$ contacts the imaginary axis are equidistant from the origin. Then $S(-\omega A - \alpha I,\epsilon)$ is clearly the reflection of $S(\omega A + \alpha I,\epsilon)$ in the origin, as shown in Figure 2. This violates Conjecture 2, since the convex hull of the spectrum of $\omega A + \alpha I$ and $-\omega A - \alpha I$ is a line segment through the origin passing between the single eigenvalue λ of $\omega A+\alpha I$ and $-\lambda$, and the minimizing λ in the definition of sep_λ must lie at one of the two points of contact on the imaginary axis. Therefore the truth of Conjecture 2 would imply the truth of Conjecture 3.

Finally, we present a counterexample to Conjecture 3, which is therefore also a counterexample to Conjectures 1 and 2. Let

$$A = \begin{bmatrix} -1 & -B & -B^2 \\ 0 & -1 & -B \\ 0 & 0 & -1 \end{bmatrix}$$

where $B >> 1$. A contour plot of $\log_{10}(\sigma_{\min}(A-\lambda I))$ in the λ plane (shapes of $S(A,\epsilon)$ for various ϵ) is shown in Figure 3 (for $B=100$); the nonconvexity of the contours is apparent. In fact, some of the $S(A,\epsilon)$ are not even simply connected! From Figure 3, we see that 0 is a local maximum of the function $\sigma_{\min}(A-\lambda)$. Thus, for example, $S(A,10^{-3})$ is essentially a disk with a small hole near the origin. In other words one can make $A+\delta A$ have any eigenvalue in an annulus about 0 with smaller $\|\delta A\|$ than is needed to make $A+\delta A$ have eigenvalue 0.

To see how much Conjecture 1 can be violated, consider the function $\sigma_{\min}(A-i\mu I)$, where μ is real. A plot of $\log_{10}(\sigma_{\min}(A-i\mu I))$ versus μ is shown in Figure 4 for $B=100$. We will show that for $\mu=2^{-1/2}$ $\sigma_{\min}(A-i\mu I)$ is at most $3^{3/2} / (2B^2)$ whereas $\sigma_{\min}(A)$ is of order $1/B$, which is much larger. Therefore $S(A,\epsilon)$ would touch the imaginary axis at about $\pm i2^{-1/2}$ for $\epsilon = O(B^{-2})$ but not the origin until $\epsilon \approx B^{-1}$.

The proof is a simple computation.

$$\sigma_{\min}(A-\lambda I) = \|(A-\lambda I)^{-1}\|^{-1} = \left\| \begin{bmatrix} \frac{-1}{1+\lambda} & \frac{B}{(1+\lambda)^2} & \frac{\lambda B^2}{(1+\lambda)^3} \\ 0 & \frac{-1}{1+\lambda} & \frac{B}{(1+\lambda)^2} \\ 0 & 0 & \frac{-1}{1+\lambda} \end{bmatrix} \right\|^{-1} .$$

When $\lambda = i\mu = 0$,

$$\sigma_{\min}(A - i\mu I) = \|A^{-1}\|^{-1} = \left\| \begin{bmatrix} -1 & B & 0 \\ 0 & -1 & B \\ 0 & 0 & -1 \end{bmatrix} \right\|^{-1} \approx \frac{1}{B}$$

for $B >> 1$. When $\lambda = i\mu \neq 0$,

$$\sigma_{\min}(A - i\mu I) = \|(A - i\mu I)^{-1}\|^{-1} \leq \left|\frac{-i\mu B^2}{(1+i\mu)^3}\right|^{-1} = \frac{(1+\mu^2)^{3/2}}{|\mu| B^2}$$

which as a function of μ reaches its minimum $3^{3/2} / (2\, B^2)$ at $\mu = 2^{-1/2}$.

If we let A be n by n and of the same structure as before:

$$A = \begin{bmatrix} -1 & -B & \cdot & \cdot & -B^{n-1} \\ & -1 & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot \\ & & & -1 & -B \\ & & & & -1 \end{bmatrix}$$

then $\sigma_{\min}(A) \approx B^{-1}$ as before and $\sigma_{\min}(A - i\mu I)$ achieves its minimum $O(B^{1-n})$ for $\mu = O(1)$. Thus for large B and/or large n, using Conjectures 1 and 2 as computational heuristics can lead to very bad results.

Note however that the matrix A is quite special: not only is it defective but it is nearly derogatory. Of course perturbing A slightly would yield a matrix with distinct eigenvalues with similarly shaped $S(A, \epsilon)$, so defectiveness per se is not essential, but nearness to a defective and derogatory matrix. It appears that if A is far from a derogatory matrix (i.e. A can be block diagonalized with one block per eigenvalue using a well-conditioned similarity), then one cannot go too far wrong using Conjecture 1 as a heuristic, and since derogatory matrices are quite rare (in the sense that a random matrix is unlikely to be very close to one [Demmel]), the heuristic is likely to be reliable.

References

[Demmel] J. Demmel, "A Numerical Analyst's Jordan Form," Dissertation, May 1983, Computer Science Dept., University of California, Berkeley

[Van Loan] C. Van Loan, "How Near is a Stable Matrix to an Unstable Matrix?" in *Linear Algebra and its Role in Systems Theory,* vol. 47 of Contemporary Mathematics, American Mathematical Society, 1985

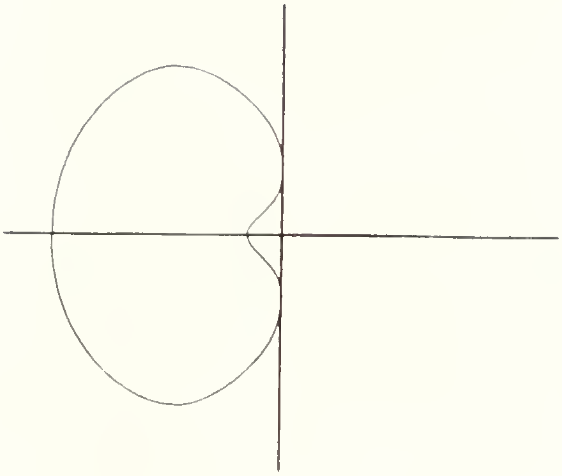


Figure 1.

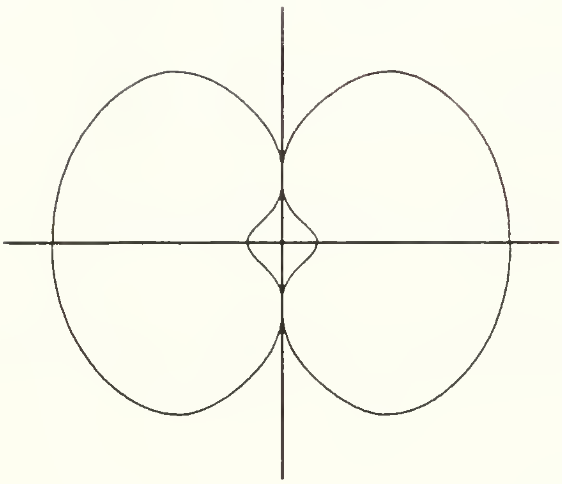


Figure 2.

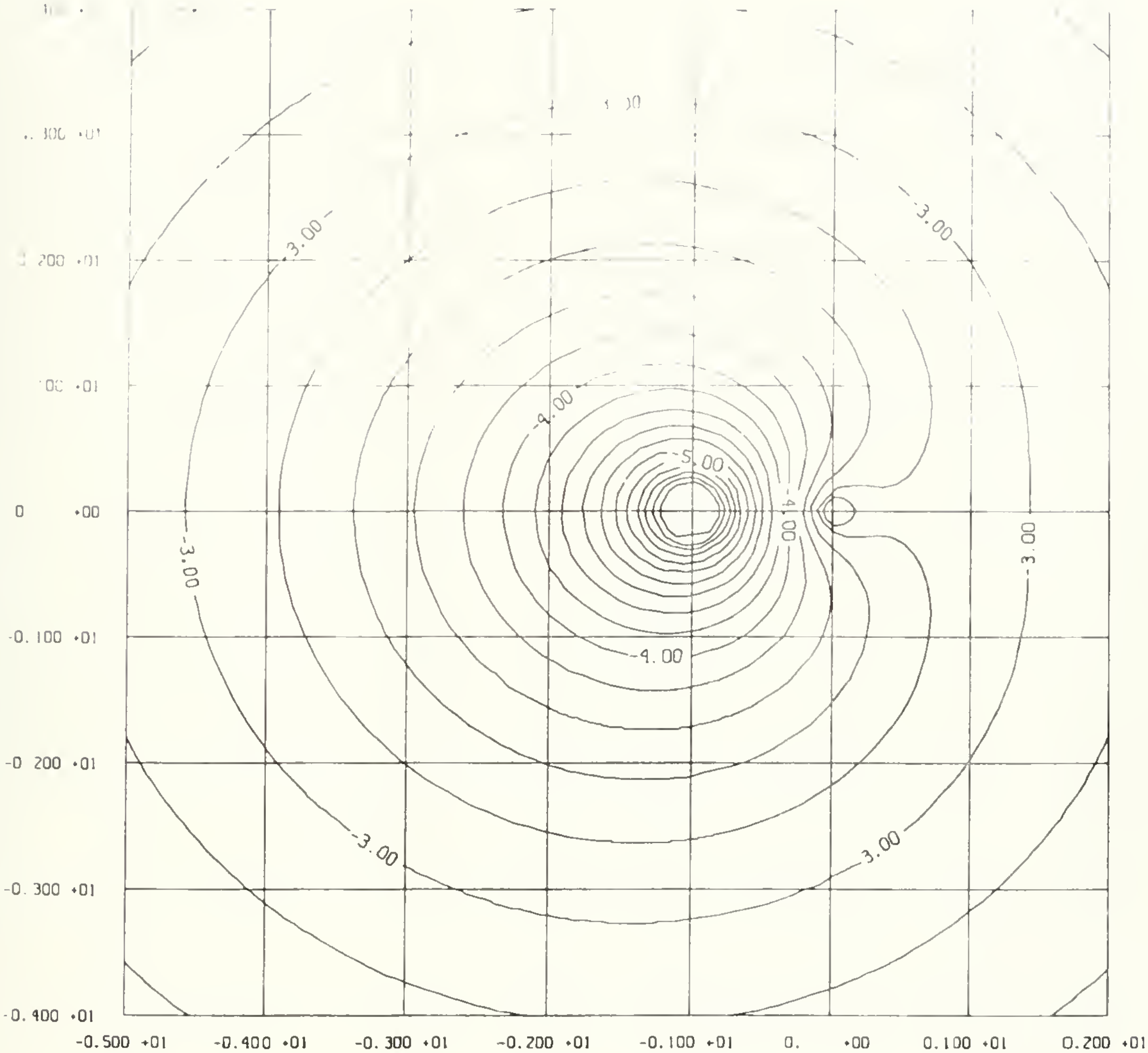


Figure 3. Contour plot of $\log_{10}(\sigma_{\min}(A - \lambda I))$ in the λ-plane

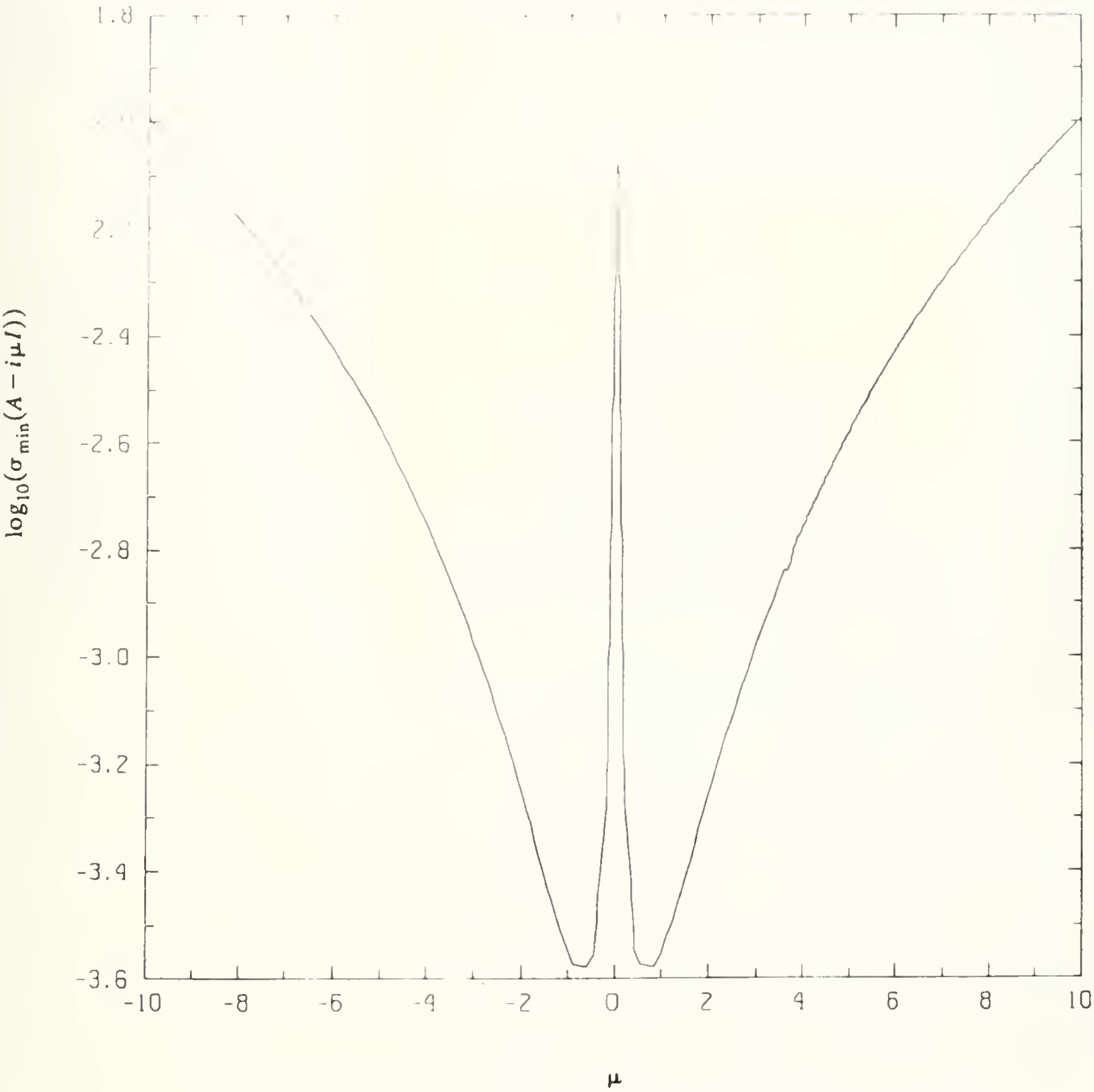


Figure 4. Graph of $\log_{10}(\sigma_{\min}(A - i\mu I))$ versus μ

NYU COMPSCI TR-220 c.1
Demmel, James Weldon
On two conjectures
regarding eigenvalue
perturbations and a common